

Annette Klosa

Einleitung

Mit diesem Band legt das Projekt *elexiko* nach fünf Jahren praktischer Artikelarbeit einen Bericht über die Erfahrungen bei der Erarbeitung eines (im Aufbau befindlichen) Online-Wörterbuches auf der Basis eines umfangreichen Korpus vor. Aus diesem Einblick in die Praxis der Erarbeitung von *elexiko* lässt sich erkennen, was es bedeutet, ein Online-Wörterbuch kontinuierlich aufzubauen, ggf. umzustrukturieren und sukzessive zu publizieren. Ein von Grund auf neu und auf der Basis eines Korpus erarbeitetes, im Internet erscheinendes Wörterbuch zum Gegenwartsdeutschen, das redaktionell erstellte und (teil-)automatisch gewonnene Angaben verbindet und sich an muttersprachliche wie nicht-muttersprachliche Nutzer, an Linguisten und Nicht-Linguisten wendet,¹ gibt es im deutschen Sprachraum sonst nicht. Vergleichbare Projekte laufen derzeit nur für wenige andere Sprachen (z.B. das *Algemeen Nederlands Woordenboek*, das am „Instituut voor Nederlandse Lexikologie“ in Leiden erarbeitet wird). Insofern sind bislang auch kaum Erfahrungen mit dieser Art von Wörterbucharbeit beschrieben und veröffentlicht worden, sodass die hier vorgelegten Berichte sowohl für Metalexikografen wie für Lexikografen aus anderen Projekten interessant sein dürften.

Die konzeptionellen Überlegungen, die im Band „Grundfragen der elektronischen Lexikographie. *elexiko* – das Online-Informationssystem zum deutschen Wortschatz“ (Haß (Hg.) 2005) dargelegt sind, werden hier, wo nötig, zum Verständnis kurz erläutert und zusammengefasst, sie sollen aber in diesem Band nicht fortgeführt werden. Änderungen an der Konzeption, die seit 2005 realisiert wurden, sind im Band „Lexikografische Portale im Internet“ (Klosa (Hg.) 2008) beschrieben. Vielmehr werden jetzt der ursprünglich zugrunde gelegten Theorie Überlegungen zur Seite gestellt, die auf den gemachten praktischen Erfahrungen basieren. So werden Dynamik und Fortentwicklung von *elexiko* dokumentiert, wobei gegebenenfalls auch auf die Praxis in anderen Wörterbüchern verwiesen wird (ohne dass hierbei Vollständigkeit angestrebt ist). Zugleich werden externe und projektbezogene Forschungsergebnisse, die seit Erscheinen der von Ulrike Haß und Annette Klosa herausgegebenen Bände publiziert wurden, berücksichtigt, auf eine vollständige Zusammenfassung der metalexikografischen Forschung wird aber verzichtet, weil die Darstel-

¹ Zu den angedachten Nutzergruppen von *elexiko* vgl. Haß (2005, S. 2f.).

lung hier schwerpunktmäßig einen Einblick in die Erfahrungen bei der praktischen Wörterbucharbeit geben soll. Es wird also insbesondere herausgearbeitet, wie die Arbeit mit dem Korpus auf die Konzeption zurückgewirkt hat und wie die einzelnen Angaben in den redaktionell bearbeiteten Wortartikeln miteinander vernetzt sind. Außerdem werden neue Überlegungen zur Datenmodellierung und Lemmatisierung angesprochen, mit denen sich das Projekt in den letzten Jahren beschäftigt hat (vgl. unten Abschnitt 1). Auch hierdurch wird Einblick in die vielfach durch die Konzeption und das Publikationsmedium Internet bedingten neuartigen Arbeitsweisen von *elexiko* gewährt.

Neben dem Verfassen von Wortartikeln und der begleitenden lexikologisch-lexikografischen Forschungstätigkeit² hat sich das Projekt schließlich mit vielen Fragen der praktischen Arbeit befasst, die in dieser Einleitung kurz beleuchtet werden: Fragen der Arbeitsgrundlagen (vgl. Abschnitt 2), methodische Fragen (vgl. Abschnitt 3) und Fragen des Online-Auftritts (vgl. Abschnitt 4).

1. Lexikografische Angaben, Datenmodellierung und Lemmatisierung

Jeder lexikografische Angabebereich in den *elexiko*-Wortartikeln ist mit eigenen Beiträgen in diesem Band vertreten, die im Folgenden in der Reihenfolge der Angaben in den Online-Artikelansichten angeordnet sind. Den Beginn macht die Untersuchung von Antje Töpel zu den **Bedeutungserläuterungen** (semantischen Paraphrasen) in *elexiko*. Die Paraphrase fasst die wichtigsten Informationen der anderen Angabebereiche zusammen und bietet dem Nutzer eine kurze Übersicht über die Bedeutung und Verwendung des jeweiligen Stichwortes. Dadurch ist die Bedeutungserläuterung besonders intensiv mit den weiteren Angabebereichen der Wortartikel vernetzt, was der Beitrag ausführlich darstellt. Weiterhin wird gezeigt, dass sich die theoretische Konzeption zur semantischen Paraphrase in *elexiko* in der redaktionellen Praxis bewährt hat und durch die beteiligten Lexikografen positiv beurteilt wird.

Annette Klosa und Petra Storjohann rekapitulieren in ihrem Beitrag zu der **semantischen Umgebung und den lexikalischen Mitspielern** die praktischen Erfahrungen bei der Ermittlung und Interpretation kontextueller Mit-

² Einen Überblick über die Vortrags- und Publikationstätigkeit als Ergebnis der lexikologisch-lexikografischen Forschung in *elexiko* geben die *elexiko*-Projektseiten im Internet. Das Projekt hat in den vergangenen Jahren außerdem zwei internationale Kolloquien organisiert: das Kolloquium „Das *elexiko*-Portal – Präsentation und Diskussion“ (10.-11. Mai 2007, Organisation: Annette Klosa) und das Kolloquium „Lexical-semantic relations from theoretical and practical perspectives“ (5.-6. Juni 2008, Organisation: Petra Storjohann).

spieler aus dem Korpus. Die Mitspieler, d. h. Satzglieder, die auch Argumentfunktion haben, flankieren einerseits die Bedeutungserläuterung, darüber hinaus konstituieren sie andererseits eine andere Art von semantischem Netz, als dies die ebenfalls im Wortartikel erfassten paradigmatischen Partnerwörter eines Stichwortes tun. Zum Teil stehen diese Angaben auch in Zusammenhang mit Informationen im Angabebereich „Besonderheiten des Gebrauchs“. In diesem Beitrag wird daneben diskutiert, wie sich das ursprünglich zugrunde gelegte framebasierte Konzept von Slots und Fillern bewährt hat. Außerdem wird an konkreten Beispielen der semantische, aber auch der enzyklopädische, assoziative und diskursive Informationsgehalt dieser Angaben gezeigt. Einige Überlegungen zu neuen Möglichkeiten der Präsentation der Mitspielerangaben sowie zur Frage, inwiefern in *lexiko* mit diesen Angaben im Ansatz Informationen eines Kollokationswörterbuches vorliegen, runden den Beitrag ab.

Der Angabebereich der **typischen Verwendungsmuster** wurde gegenüber der ursprünglichen Konzeption deutlich verändert: Die Klassifizierung und Sortierung der Muster erfolgt im Vergleich zum Ausgangskonzept anhand syntaktischer Kriterien, außerdem wurde für die so genannten Klammerelemente ein Kategorieninventar entwickelt. Zudem wurden Überlegungen zur Illustration variabler Elemente innerhalb der Klammern angestellt, damit diese leichter von invariablen Elementen zu unterscheiden sind. Christine Möhrs stellt in ihrem Beitrag die jetzige Extrahierung und Struktur der typischen Verwendungsmuster in *lexiko* dar und bewertet sie. Sie beleuchtet daneben die inhaltliche Vernetzung innerhalb eines redaktionell bearbeiteten Wortartikels zwischen den typischen Verwendungen und anderen Angabebereichen. Diese Vernetzung steigert insofern die Qualität eines Wortartikels, als ein Benutzer bei der Rezeption über die Angabebereiche hinweg bestimmte Aspekte wiederfindet.

Petra Storjohann erläutert in ihrem Beitrag zu den **sinn- und sachverwandten Wörtern**, wie die Arbeit mit dem *lexiko*-Korpus in Verbindung mit lexikologischer Forschung dazu geführt hat, dass gegenüber der Ursprungskonzeption nun paradigmatische Relationen in *lexiko* beschrieben werden, die bisher nicht lexikologisch beschrieben und lexikografisch erfasst worden sind. Wo genau die Vor- und Nachteile der Arbeit mit umfangreichen elektronischen Textquellen liegen und wie lange die Bearbeitung dieses Informationsbereiches dauert, sind nur einige der Aspekte, die hier erörtert werden. Es erfolgt auch eine kritische lexikologische Auseinandersetzung mit den bestehenden Klassifikationen, die für die Einordnung der paradigmatischen Strukturen als Grundlage gewählt wurden. Dieser Beitrag thematisiert anhand vieler Beispiele

auch, wo die Diskrepanz zwischen Sprache als System und Sprache im Gebrauch am stärksten erkennbar wird, und erläutert den derzeitigen Umgang mit der Anwendung bestehender klassifikatorischer Modelle.

Die Angaben zu den Besonderheiten des Gebrauchs enthalten inzwischen auch die Angabe zur **themengebundenen Verwendung des Stichwortes** bzw. die Darstellung der Zugehörigkeit eines Stichwortes zu einem bestimmten Diskursvokabular. Ulrich Schnörch stellt diesen Angabebereich in seinem Beitrag vor und widmet sich dessen Inhalt, Genese und seiner konzeptionellen wie methodischen Vernetzung mit anderen Angabetypen anhand vieler einschlägiger Beispiele.

Annette Klosa gibt in ihrem Beitrag zu **grammatischen Angaben und Angaben zur Wortbildung** in *lexiko* einen kurzen Überblick über die Änderungen und Ergänzungen, die zu diesen Angabebereichen in den vergangenen Jahren vorgenommen wurden. Hierbei wird besonders deutlich, wie sich die Arbeit mit einem sehr großen Korpus auf diese Bereiche ausgewirkt hat. Sie verweist außerdem auf die Vernetzung mit anderen Angaben im Wortartikel und schlägt erweiterte Suchmöglichkeiten für diese Angabebereiche vor.

Ein abschließender Beitrag von Annette Klosa fasst weitere, **kleinere Neuerungen und Ergänzungen** zu einzelnen Angabebereichen (z. B. Frequenzangabe, Illustrationen) und Fragen der Lemmatisierung (z. B. von Pronomen) zusammen. Es werden auch Punkte angesprochen, deren Konzeption noch offen ist (z. B. Lemmatisierung von Wortbildungselementen). Hieraus wird deutlich, dass *lexiko* in vielerlei Hinsicht ständigen Entwicklungen unterworfen ist.

Neben den lexikografischen Angaben gibt es auch im Bereich der **Datenmodellierung** Interessantes aus fünf Jahren lexikografischer Praxis zu berichten: In *lexiko* werden die lexikografischen Wortartikel in einer granularen, maßgeschneiderten XML-Struktur erarbeitet. Der Beitrag von Carolin Müller-Spitzer legt die praktischen Erfahrungen dar, die alle Beteiligten im Team mit dieser Art der Modellierung gemacht haben. Im Fokus steht dabei diese Praxis, daneben werden aber auch Aspekte wie die Umarbeitung bereits geschriebener Wortartikel nach konzeptionellen inhaltlichen Änderungen oder Erkenntnisse aus der Weiterverarbeitung der Daten für die Darstellung sowie für die Entwicklung von Suchfunktionen ausgeführt. Es wird gezeigt, dass das Konzept für die Modellierung insgesamt erfolgreich im lexikografischen Prozess Anwendung gefunden hat, dass sich aber bestimmte Rahmenbedingungen (wie die mangelnde personelle Ausstattung für die technische Unterstützung) als

problematisch erwiesen haben. Insgesamt soll dieser Beitrag damit die lexikografische Praxis im Vergleich mit der theoretischen Konzeption der Modellierung beleuchten.

Schließlich verändert sich auch die *elexiko*-Stichwortliste. Eine wichtige Lücke in der bisherigen Konzeption konnte inzwischen durch die Entscheidung zur Lemmatisierung und **lexikografischen Behandlung von Eigennamen** in *elexiko* geschlossen werden, die im Beitrag von Annette Klosa und Sabine Schoolaert vorgestellt wird. Vor dem Hintergrund der Diskussion möglicher Definitionen von „Eigennamen“ und der gängigen lexikografischen Behandlung von Namen diskutiert dieser Beitrag, ob in *elexiko* neben Gattungsprädikaten auch Eigennamen lemmatisiert werden sollen. Außerdem beschreibt der Beitrag die Behandlung von Eigennamen in *elexiko* hinsichtlich ihrer Erfassung, Klassifizierung und Darstellung und erläutert unterschiedliche Angabetypen. Ein Ausblick auf Suchoptionen zu den Eigennamen schließt die Überlegungen ab.

2. Die Arbeitsgrundlagen von *elexiko*

Zu den Arbeitsgrundlagen von *elexiko* gehören u. a. das *elexiko*-Korpus, die redaktionelle Arbeitsumgebung und die DTD-Dokumentation mit integriertem Redaktionshandbuch. In allen diesen Bereichen hat sich seit Projektbeginn einiges verändert, worauf an dieser Stelle kurz eingegangen werden soll.

Das ***elexiko*-Korpus** konnte seit Projektbeginn stark erweitert werden:³ Umfasste es anfangs knapp 1,3 Milliarden Textwörter aus 19 Quellen (vgl. Storzjohann 2005b, S. 67), enthält es derzeit (Stand 2010) 2,8 Milliarden Textwörter aus 32 Quellen. Es wird auch weiterhin als virtuelles Korpus aus DEREKO,⁴ dem deutschen Referenzkorpus des IDS, zusammengestellt und umfasst ausschließlich Zeitungs- und Zeitschriftenquellen.

Dank der Erweiterungen gelang es, den Anteil an Quellen aus der Schweiz und aus Österreich besser auf die tatsächlichen Sprecherzahlen in diesen deutschen Sprachräumen abzustimmen. Inzwischen sind auch West-, Mittel- und Norddeutschland durch die Aufnahme entsprechender Quellen angemessener ver-

³ Für seine Bemühungen um den Ausbau der IDS-Korpora insgesamt, von dem *elexiko* direkt profitiert, ist *elexiko* dem Projekt „Ausbau und Pflege der Korpora geschriebener Gegenwartssprache: Das deutsche Referenzkorpus DEREKO“ zu großem Dank verpflichtet, für die ständige Pflege und Erweiterung des *elexiko*-Korpus dankt *elexiko* dem Projekt COSMAS II am IDS.

⁴ Zum Deutschen Referenzkorpus (DEREKO) des IDS vgl. www.ids-mannheim.de/kl/projekte/korpora/ (Stand: 07.05.2010).

treten. Erfreulicherweise konnte der Anteil der öffentlich zugänglichen Quellen insgesamt erhöht werden, sodass *elexiko* dem Ziel, wenigstens einen großen Teil des Wörterbuchkorpus der Öffentlichkeit durch direkte Verlinkung von Wortartikeln ins Korpus zugänglich machen zu können, näher gekommen ist.

Das *elexiko*-Korpus strebt an, „die Grundgesamtheit der deutschen standard-sprachlichen Gemeinsprache in angemessener Weise zu spiegeln“ (Storjohann 2005b, S. 58), wobei unter Gemeinsprache der Teil einer Sprache verstanden wird, den die Mehrheit der Sprecher einer Sprachgemeinschaft teilen. Um diesen Ausschnitt des Deutschen zu beschreiben, eignen sich überregionale Zeitungs- und Zeitschriftentexte sehr gut, weil sie an eine Sprachgemeinschaft insgesamt adressiert und daher stark von standardisierter Sprache geprägt sind. Daneben ist die Analyse der Sprache regionaler Zeitungen sinnvoll, da so regionale Sprachunterschiede sowie Unterschiede zwischen standardisierter Gemeinsprache und regionaler Sprache deutlich werden. Für die Ziele von *elexiko*, den öffentlichen Sprachgebrauch zu untersuchen (vgl. hierzu Abschnitt 3) und die standardsprachliche Gemeinsprache zu dokumentieren, eignet sich ein (z. B. regional und zeitlich ausgewogenes) Zeitungs-/Zeitschriftenkorpus von sehr großem Umfang besonders gut.⁵

Die **redaktionelle Arbeitsumgebung** von *elexiko* besteht seit Projektbeginn aus den folgenden Kernstücken:

- COSMAS-II-Schnittstelle⁶ zum *elexiko*-Korpus, mithilfe derer die lexikografischen Daten gewonnen werden;
- XML-Editor, in dem die lexikografischen Daten ediert werden;
- Schnittstelle zur ORACLE-Datenbank, in der die Daten gespeichert werden, in der sie durchsuchbar sind und aus der sie für die Online-Präsentation exportiert werden.

Inzwischen wurde durch die Arbeit des Projektes „Texttechnologie am IDS“ der Zugriff auf die Datenbank über das „Lexikografische Redaktions- und Recherchesystem für digitale Wörterbücher (EDAS = Electronic Dictionary Administration System⁷)“ wesentlich verbessert. So können nun die einzelnen

⁵ Für andere Untersuchungszwecke eignet sich ein Korpus wie das *elexiko*-Korpus dagegen weniger. Man sollte außerdem nicht generell davon ausgehen, dass Zeitungstexte die deutsche Sprache insgesamt sehr gut repräsentieren.

⁶ Zu COSMAS II („Corpus Search, Management and Analysis System“) vgl. <https://cosmas2.ids-mannheim.de/cosmas2-web/> (Stand: 07.05.2010).

⁷ Vgl. hierzu <http://www.ids-mannheim.de/gra/texttechnologie.html> (Stand: 07.05.2010) und Müller-Spitzer/Schneider (2009). Zu weiteren Einzelheiten vgl. Abschnitt 7 im Beitrag „Der Einsatz einer maßgeschneiderten, feingranularen XML-Modellierung im lexikografischen Prozess“ von Carolin Müller-Spitzer in diesem Band.

redaktionellen Arbeitsschritte und Zuständigkeiten mit einer optimierten Rechteverwaltung genauer abgebildet werden. Viel wichtiger ist aber die Möglichkeit, in der Datenbank direkt erweiterte Suchen definieren zu können.⁸ Im redaktionellen Arbeitsalltag stellt dies eine deutliche Erleichterung dar, weil Stichwörter mit ähnlichen Merkmalen leicht gefunden werden können und so ein konsistenteres Arbeiten möglich ist. Zur redaktionellen Arbeitsumgebung ist in den letzten Jahren auch eine große Zahl komfortabel zur Verfügung stehender elektronischer Sekundärquellen, also andere Wörterbücher auf CD-ROM oder im Internet, hinzugekommen.⁹

Dem Projektteam steht inzwischen auch ein im Projekt BZ*Velexiko*¹⁰ entwickelter Verweismanager zur Verfügung,¹¹ der die Erstellung konsistenter Verweise im Wörterbuch ermöglicht und so nicht nur den Lexikografen die Arbeit erleichtert. Die Verbesserung der Konsistenz in Bezug auf Verweise steigert gleichzeitig die Nutzbarkeit von *elexiko* für die Wörterbuchbenutzer. Zugleich wurden einige Ergänzungen zum XML-Editor programmiert (z. B. eine Kommentarfunktion, eine Textbausteinfunktion), die nicht unerheblich zur Arbeitserleichterung beitragen.

Mit fortschreitender Zahl redaktionell bearbeiteter Wortartikel wurde der Wunsch nach einem **Redaktionshandbuch** in der Projektgruppe immer größer. Hier sollten die zugrunde gelegte Konzeption, alle Entscheidungen zu Formulierungen (z. B. der Satzbaupläne bei Verben), Richtlinien für die Verwendung von Hinweisen oder Kommentaren, Leitlinien für die Auswahl von Belegen etc. mit konkreten Beispielen festgehalten werden, um die Arbeit an neuen Wortartikeln sowie die Einarbeitung neuer Kollegen zu erleichtern. Im Redaktionshandbuch sollten also alle die Regeln festgehalten werden, die selbst eine sehr komplexe und granulare XML-Struktur wie in *elexiko*, die die Lexikografen sicher durch den hochkomplexen Wortartikel führt, nicht umsetzen kann.

Gleichzeitig war unumgänglich, die DTD sorgfältig zu dokumentieren, auch, um Änderungen und Ergänzungen festhalten zu können. Deshalb wurde zum Grundaufbau des Redaktionshandbuches bzw. der **DTD-Dokumentation** eine Software gewählt, die aus DTDs automatisch eine Rohfassung der Dokumentation in XML erstellt. Somit waren alle DTD-Elemente und Attribute als Felder vorgesehen, mit frei zu definierenden Feldern für Erläuterungen zur Mo-

⁸ Sehr komplexe Abfragen werden auch weiterhin in der Arbeitsstelle Zentrale DV-Dienste des IDS programmiert, wofür *elexiko* zu großem Dank verpflichtet ist.

⁹ Für die Anschaffung dieser Quellen dankt das Projekt der IDS-Bibliothek.

¹⁰ Zum Projekt BZ*Velexiko* vgl. das Vorwort in diesem Band.

¹¹ Vgl. hierzu Meyer/Müller-Spitzer (2010).

dellierung, inhaltlichen lexikografischen Richtlinien etc. Im Laufe der Zeit wurde diese DTD-Dokumentation zu einem echten Redaktionshandbuch ausgebaut.¹² Diese XML-Version der Dokumentation kann in ein HTML- und ein RTF-Format exportiert werden. Im Projekt wird ausschließlich das HTML-Format genutzt, da das Redaktionshandbuch kontinuierlich ausgebaut wird und im Ausdruck etwa 500 Seiten umfassen würde. Es kann beim Schreiben der Wortartikel aus dem XML-Editor heraus komfortabel und kontextsensitiv aufgerufen werden.

3. Methodik in *elexiko*

elexiko wird grundsätzlich **korpusgestützt** erarbeitet, d. h. alle Erkenntnisse, die in die Angabebereiche einfließen, werden aus dem zugrunde gelegten Wörterbuchkorpus gewonnen: In *elexiko* gibt es keine lexikografische Angabe, die nicht im *elexiko*-Korpus belegt werden kann, es werden keine Angaben aus anderen Wörterbüchern einfach übernommen. Falls nötig, wird explizit darauf verwiesen, dass bestimmte Phänomene (entgegen der Angabe in anderen Wörterbüchern) im Korpus nicht gefunden werden konnten. Dies ist z. B. bei in Sekundärquellen gebuchten, im *elexiko*-Korpus aber nicht nachzuweisenden Lesarten (= Einzelbedeutungen) der Fall. In solchen seltenen (und oft fachsprachlichen) Fällen wird der Nutzer in Hinweisen über die entsprechenden Wörterbuch- und Korpusbefunde informiert. Diese wichtigste Methode für die Erarbeitung von *elexiko*, das Prinzip der **Korpusbasiertheit** (vgl. Haß 2005, S. 7ff.), hat sich, wie die Beiträge zu den einzelnen Angabebereichen in diesem Band deutlich zeigen, durchgehend bewährt.

Bei der korpusgestützten Erarbeitung von *elexiko* kommen zwei verschiedene Vorgehen zum Einsatz: die korpusgesteuerte sowie die korpusbasierte Ermittlung von Angaben. Bei der **korpusgesteuerten** Auswertung benutzt der Lexikograf das Korpus explorativ, er befragt es ohne Vorannahme mithilfe verschiedener Korpusrecherche- und -analysetools, die Ergebnisse werden dann analysiert, bewertet und in den für die Wortartikel vorgesehenen Angaben beschrieben. Bei diesem Vorgehen profitiert *elexiko* insbesondere vom Verfahren der statistischen Kollokationsanalyse (vgl. Belica 1995), das in COSMAS II implementiert ist, und den Recherchemöglichkeiten in der Kookkurrenzdatenbank CCDB (vgl. Belica 2001-2007).¹³

¹² Zu weiteren Einzelheiten sowie einer Abbildung aus dem Redaktionshandbuch vgl. Abschnitt 3 im Beitrag „Der Einsatz einer maßgeschneiderten, feingranularen XML-Modellierung im lexikografischen Prozess“ von Carolin Müller-Spitzer in diesem Band.

¹³ Für die kollegiale Zusammenarbeit zum Einsatz der statistischen Kollokationsanalyse wie der CCDB in *elexiko* ist *elexiko* dem Projekt „Methoden der Korpusanalyse und -erschlie-

Wird das Korpus **korpusbasiert** ausgewertet, geht ein Lexikograf von einer bestimmten Annahme zu einem sprachlichen Phänomen aus (z. B. davon, dass zwei Lexeme in bestimmten Lesarten zueinander in antonymischer Relation stehen) und sucht im Korpus gezielt nach Belegen dafür. Beide Methoden werden in *ellexiko* in Kombination angewandt (vgl. hierzu Storjohann 2005a und Klosa 2007), was sich insbesondere in Hinblick auf den Angabebereich der sinn- und sachverwandten Wörter, aber auch sonst sehr bewährt hat. In Abhängigkeit vom Analysezweck wird entweder zunächst der korpusgesteuerten Analyse des Korpus der Vorzug gegeben und eine korpusbasierte Auswertung erfolgt erst in einem zweiten Schritt, oder beide Methoden werden parallel angewendet.

Da *ellexiko* erst noch aufgebaut und nicht vollständig zu einem bestimmten Zeitpunkt publiziert wird, erscheint es nicht sinnvoll, die Wortartikel alphabetisch von A bis Z, sondern in anders zu definierenden Teilschritten zu erarbeiten. Nach Abschluss der Arbeit am *ellexiko*-Demonstrationswortschatz¹⁴ war es deshalb nötig, Kriterien für die Auswahl von zu bearbeitenden Teilwortschätzen (so genannten Modulen) zu finden, zu prüfen und für den weiteren Projektverlauf nutzbar zu machen (vgl. Haß 2005, S. 14). Als Hauptkriterium für die Auswahl weiterer **Bearbeitungsteilwortschätze** wurde daraufhin die Zugehörigkeit von Stichwörtern zu einzelnen Frequenzschichten¹⁵ herangezogen. Auf dieser Grundlage sind zwei Bearbeitungsteilwortschätze definiert worden: das *Lexikon zum öffentlichen Sprachgebrauch* und die Menge an niedrigfrequenten Stichwörtern, die mit automatisch generierten Angaben versehen wird.

Das *Lexikon zum öffentlichen Sprachgebrauch* ist das erste Modul nach dem Demonstrationswortschatz, das als Bearbeitungsteilwortschatz festgelegt wurde. Dieser Wortschatz deckt sowohl Themen aus Politik und Gesellschaft als auch speziellere Sachverhalte ab. Gut die Hälfte der insgesamt rund 2.700 hochfrequenten Wörter (jeweils zwischen 10.000- und 500.000-mal im *ellexiko*-Korpus belegt) sind Nomen, die häufig in die zentralen politischen und gesellschaftlichen Diskurse, wie sie im *ellexiko*-Korpus erscheinen, eingebettet sind. Sie werden hauptsächlich durch Verben und Adjektive ergänzt, die zu einem

Bung“ (insbesondere dem Schwerpunkt „Kookkurrenzanalyse und deren Erschließung“) zu großem Dank verpflichtet.

¹⁴ Zu Informationen über den *ellexiko*-Demonstrationswortschatz vgl. Haß (2005, S. 15) und die *ellexiko*-Projektseiten im Internet.

¹⁵ Vgl. hierzu Abschnitt 1.1 im Beitrag „Von Abbildung bis Worтеlement: Weitere Ergänzungen und Änderungen in *ellexiko*“ von Annette Klosa in diesem Band.

geringeren Teil selbst diskursgebunden sind (z. B. *reformieren*, *global*, *sozial-verträglich*), die aber auch zur Versprachlichung der Diskurse benötigt werden (z. B. *feststellen*, *abstimmen*). Ein zeitungs- und zeitschriftensprachliches Korpus wie das *elexiko*-Korpus (vgl. Abschnitt 2) eignet sich besonders gut als Beschreibungsgrundlage für diesen Wortschatzausschnitt.

Beim *Lexikon zum öffentlichen Sprachgebrauch* handelt es sich nicht um ein Wörterbuch zum Fach- oder Spezialwortschatz der Politik, sondern um ein Wörterbuch, in dem der allgemein geläufige Wortschatz der öffentlichen Diskussion um politische und gesellschaftliche Themen semantisch und pragmatisch ausführlich, fein differenziert und sprachreflektierend dargestellt wird. Etwa die Hälfte der Stichwörter ist inzwischen bearbeitet, und die in den Beiträgen dieses Bandes geschilderte Praxis bezieht sich überwiegend auf die Bearbeitung dieser Stichwörter.

Niedrigfrequente Stichwörter sind in *elexiko* Wörter, die weniger als 500-mal im *elexiko*-Korpus belegt sind. Neben der Erarbeitung des *Lexikons zum öffentlichen Sprachgebrauch* wird *elexiko* derzeit durch **(teil-)automatisch ermittelte Angaben** für diese niedrigfrequenten Stichwörter gefüllt. So werden in diesem *elexiko*-Modul zusätzlich zu den orthografischen Angaben automatisch ermittelte Belege aus dem *elexiko*-Korpus ergänzt, wie die Angaben zur Beleglage des Stichwortes im *elexiko*-Korpus (vgl. das Beispiel *glücklich* in Abb. 1).¹⁶ Grammatische Angaben zur Wortart und (bei Wörtern, die durch Wortbildung entstanden sind) zu den Bestandteilen, aus denen die Bildung besteht (z. B. bei Ableitungen die Basis und das Affix), sollen das Angebot zukünftig ergänzen.¹⁷ Daneben wird an der automatischen Ermittlung von Wortbildungsprodukten zu einem Stichwort gearbeitet.¹⁸

Da für *elexiko* eine möglichst gute, d. h. eine wissenschaftlichen wie lexikografischen Ansprüchen gerecht werdende Qualität der Angaben angestrebt ist, wird, außer bei den Belegen, auf die Präsentation rein automatisch ermittelter Angaben verzichtet. Die automatisch ermittelten Belege werden als solche gekennzeichnet (vgl. Abb. 1).

¹⁶ Damit sich Nutzer nicht darüber wundern, dass bei manchen Stichwörtern automatisch ausgewählte Belege erscheinen, bei vielen anderen aber nicht, wurden inzwischen auch andere, durchschnittlich oder hochfrequente Stichwörter mit Belegen versehen.

¹⁷ An der Überprüfung automatisch ermittelter Angaben zur Wortbildung der Stichwörter wird derzeit gearbeitet; diese Analysen wurden mithilfe des Tools Morphisto erstellt (vgl. hierzu www.ids-mannheim.de/ll/TextGrid/morphisto.html, Stand: 10.05.2010).

¹⁸ Vgl. hierzu Abschnitt 2.2 im Beitrag „Von Abbildung bis Wortelement: Weitere Ergänzungen und Änderungen in *elexiko*“ von Annette Klosa in diesem Band.

glücklich

Orthografie

Normgerechte Schreibung: glücklich
Worttrennung: glück|lich

Belege (automatisch ausgewählt)

Egal. Was soll's. Wir werden uns wegen dieser kleinen Zeitumstellung doch nicht aus dem Gleichgewicht bringen lassen. Wäre ja gelacht. Sind wir doch seit dem neuen Clowndoctor-Film "Patch Adams" allesamt exzessiv **glücklich**. Jawohl, immerhin ist jetzt auch schon Hollywood draufgekommen, daß Lachen gesund ist, Kinosäle und -kassen füllt. (P99/MAR.12395 Die Presse, 29.03.1999, Ressort: up; Der Montag mit Elisabeth Krammer)

Seit 29 Jahren, beteuert das ehemalige Mitglied der legendären Monty-Python-Truppe, lebt er bereits in London, und da sei es doch ganz normal, einen britischen Paß zu kriegen. Er durfte die amerikanische Staatsbürgerschaft trotzdem behalten: "Was ich heute aber als Heimat betrachte ist ganz klar. Für '12 Monkeys' war ich sieben Monate auswärts. Sie können sich gar nicht vorstellen, wie **glücklich** ich war, wieder in London zu sein. Man geht durch die Straßen, und die Menschen sehen ganz verschieden aus." (X96/MAR.03078 Oberösterreichische Nachrichten, 19.03.1996, Ressort: Kultur; Der kahrlasierte Schädel bringt's)

Demokratie kann totalitär sein, wenn die Herrschaft der Mehrheit rücksichtslos die Rechte der Minderheiten verletzt und ihre Chance beeinträchtigt, Mehrheit zu werden. Liberalismus ist eine Auffassung vom Herrschaftsgrad. da der Liberale weiß, daß in jeder Gesellschaft das Element der Macht wirkt und diese Macht nicht zu eliminieren ist, versucht er sich nicht an der Abschaffung, sondern sieht seine Aufgabe in der Begrenzung, Aufteilung und Kontrolle der Macht und im Offenhalten der Chance zur Ablösung derjenigen, welche die Macht ausüben. Liberalismus und Demokratie sind in vielen Ländern eine **glückliche** Verbindung eingegangen. die geistige Stärke des Liberalismus bedingt seine organisatorische Schwäche. (H85/QK1.00000 Flach, Karl-Hermann: Noch eine Chance für die Liberalen Oder: Die Zukunft der Freiheit. - Frankfurt a.M., 1971)

Dieses Stichwort gehört im *lexiko*-Korpus der Frequenzschicht IX (10.001-50.000 mal belegt) an. Es ist in 18 verschiedenen Zeitungen oder Zeitschriften aus 26 Jahrgängen belegt.

Weitere Informationen:

Automatisch ermitteltes Kookkurrenzprofil von **glücklich** in der [CCDB](#)
Grammatische Informationen (z.B. Angabe der Wortart, Flexionstabellen) unter [canoo.net](#).

Abb. 1: Beispiel für einen Wortartikel mit (teil-)automatisch generierten Angaben

So werden z. B. automatisch erstellte Wortbildungsanalysen redaktionell überprüft, bevor sie online erscheinen. Belege zu einem Stichwort werden nicht nach rein statistischen Zufallskriterien aus dem Korpus ermittelt, sondern unter Hinzunahme der folgenden, die Auswahl weiter steuernden Kriterien:¹⁹

- Die Belege müssen mindestens aus drei verschiedenen Quellen und Jahrgängen stammen.
- Der Belegumfang soll maximal drei Sätze vor dem Satz mit dem Stichwort und einen Satz danach umfassen, höchstens aber einen Absatz.
- Belege, in denen das Stichwort z. B. in einer Überschrift steht, werden dadurch ausgeschlossen, dass eine Mindestsatzlänge (beruhend auf der durchschnittlichen Satzlänge im *lexiko*-Korpus) angenommen wird.

Dies bedeutet aber auch, dass die Füllung der Stichwörter um neue Angaben langsamer vorangeht, als dies bei einer ausschließlich automatischen Methode der Fall wäre. Dafür wird aber (durch Fehlerkorrekturen) eine bessere Qualität der Angaben gewährleistet (vgl. Klosa 2010). Übergangsweise ermöglichen Hyperlinks auf externe Online-Angebote (vgl. Abb. 1) Nutzern, grammatische Informationen zu einem Stichwort nachzuschlagen (bei canoo.net) bzw. statistisch signifikante Kollokatoren zu einem Stichwort einzusehen (in der Kookkurrenzdatenbank CCDB des IDS). Bei bearbeiteten Stichwörtern können diese Hyperlinks entfallen, da die entsprechenden Angaben im Rahmen der Wortartikelbearbeitung ermittelt wurden und im fertigen Wortartikel zu finden sind.

Da *lexiko* fortlaufend publiziert wird, können permanent Nachträge in die oder Löschungen aus der Stichwortliste sowie Korrekturen und Ergänzungen in redaktionell bearbeiteten Wortartikeln vorgenommen werden (vgl. Hahn et al. 2008, S. 80ff.). Dies hat einerseits den Vorteil, dass die Stichwortliste und die Wortartikel aktuell gehalten werden können, andererseits aber den Nachteil, dass ohne Kennzeichnung der jeweiligen Stichwortlisten- bzw. Artikelversion das Wörterbuch im wissenschaftlichen Kontext nicht zitierbar ist. Derzeit wird in *lexiko* nicht markiert, wann die Stichwortliste oder einzelne Wortartikel zuletzt bearbeitet bzw. aktualisiert wurden. Es gibt auch (noch) kein Archiv, auf das Nutzer bei der Suche nach älteren Fassungen zugreifen könnten. Die Frage nach **Archivierung und Versionierung** der *lexiko*-Substanzen steht daher dringend an. Deshalb werden die tatsächlichen Erwartungen der

¹⁹ Die automatische Ermittlung der Belege wird dankenswerterweise durch das Projekt COSMAS II am IDS übernommen. Im Fall falsch ausgewählter Belege können andere Belege durch die Lexikografen nachgetragen werden. Es wird dabei artikelintern vermerkt, dass die Belegauswahl manuell verändert wurde.

Nutzer an ein fortwährend publiziertes Online-Wörterbuch mithilfe von Benutzerbefragung und -beobachtung im Rahmen der im Projekt BZV*lexiko* durchgeführten Benutzungsforschung abgefragt. Auf der Basis dieser Erkenntnisse sollen Archivierung und Versionierung dann konzeptionell und technisch umgesetzt werden.

4. *ellexiko* online

Zu Beginn der Projektlaufzeit war *ellexiko* umfassend sowohl als korpusgestütztes, gegenwartssprachliches Bedeutungswörterbuch wie als Wortschatzinformationssystem zum Deutschen, das andere Module umfassen sollte, geplant (vgl. Haß 2005). Bedingt durch die enorme Fülle der Aufgaben im Projekt *ellexiko* sowie aufgrund verschiedener Entwicklungen in den ursprünglichen Modulprojekten „Neologismen der 90er Jahre“ und „Usuelle Wortverbindungen“ und nicht zuletzt durch die Ausweitung der Planungen auf weitere Online-Wörterbuchprojekte wurde zwischenzeitlich die Portalfunktion aus *ellexiko* ausgegliedert und dem neuen Projekt **OWID** („Online-Wortschatz-Informationssystem Deutsch“)²⁰ übertragen (vgl. Klosa 2008, S. 3; Müller-Spitzer 2008). *ellexiko* ist nun eines unter mehreren Wörterbüchern, die innerhalb von OWID publiziert werden. Allerdings ist es das von der Gesamtzahl der Stichwörter, der Menge an redaktionell bearbeiteten Wortartikeln und der Breite der Angaben her umfangreichste Nachschlagewerk in diesem Portal. *ellexiko* ist, wo es möglich und sinnvoll erscheint, mit den anderen Wörterbüchern in OWID vernetzt, ein weiterer Ausbau der Vernetzungen wird inzwischen im Projekt BZV*lexiko* angestrebt.

Neben dieser großen konzeptionellen und organisatorischen Änderung hat sich auch die **Benutzeroberfläche** von *ellexiko* gewandelt.²¹ Das erste Online-Angebot von *ellexiko* umfasste nur die Stichwortliste (2003), in einem zweiten Schritt kamen die Wortartikel des *ellexiko*-Demonstrationswortschatzes mit ersten komplexeren Recherchemöglichkeiten hinzu (2004). Mit der Freischaltung der Wortartikel des Modulprojektes „Neologismen der 90er Jahre“ erhielt das Online-Angebot von *ellexiko* ein neues Design, das die beiden Funktionen von *ellexiko* (Wörterbuch und Portal) auch in der Benutzeroberfläche sichtbar machte (2007, vgl. Bock 2008). Daneben wurde die Präsentation der *ellexiko*-Wortartikel verbessert und neue erweiterte Suchmöglichkeiten konnten freigeschaltet werden. Inzwischen wurde das Design der Wortartikel an

²⁰ Zu OWID vgl. www.owid.de (Stand: 10.05.2010).

²¹ Eindrücke von den verschiedenen Entwicklungsstufen der Benutzeroberfläche vermitteln die vielen Abbildungen in den im Laufe der Projektzeit erschienenen Publikationen.

die neue Benutzeroberfläche von OWID angepasst, wobei zugleich die Möglichkeit zu Optimierungen bestand (2010). Ein weiterer Ausbau der **Suchmöglichkeiten** ist konzeptionell vorbereitet und wird in absehbarer Zeit im Zusammenhang mit dem Ausbau portalübergreifender Rechercheangebote in OWID realisiert werden. Jeder dieser Schritte bedeutete nicht nur einen erheblichen Programmieraufwand, sondern auch, dass so genannte Stylesheets, welche die Umwandlung der XML-Inhalte für die Ansicht in HTML steuern, geschrieben und getestet wurden. Der redaktionelle Input bestand hier im Entwurf der Wortartikelansichten und der Festlegungen dazu, welche Angaben online wie erscheinen sollen, sowie in gründlichem Testen der Umsetzung dieser Vorgaben.

Das Online-Angebot von *ellexiko* wurde schließlich um umfangreiche **Wörterbuchaußentexte** ergänzt, womit einerseits metalexikografische Vorschläge zu diesem Thema realisiert wurden (vgl. Klosa 2009), andererseits vor allem aber die Benutzerfreundlichkeit von *ellexiko* deutlich verbessert werden konnte. Die ausführlichen Benutzungshinweise wenden sich vornehmlich an Nutzer ohne germanistischen Fachhintergrund. Hier werden alle Fachtermini, die in den Wortartikeln erscheinen, erläutert, wobei gegebenenfalls auch auf Einträge in *grammis*,²² dem grammatischen Informationssystem des IDS, verwiesen wird. Daneben wird die lexikografische Konzeption (z. B. Methodik, Arbeitsgrundlagen) für das Fachpublikum auf den Internetseiten des Projektes vorgestellt.

5. Schlussgedanken

Dass *ellexiko* ein im Aufbau befindliches Online-Wörterbuch ist, zeigt sich nicht nur an den in diesem Band beschriebenen Beispielen für konzeptionelle Ergänzungen, Änderungen oder Neuerungen bezüglich der behandelten Artikeltypen, der Stichwortliste, der Datenmodellierung und verschiedener lexikografischer Angaben, sondern auch in jedem einzelnen Beitrag dieses Bandes: *ellexiko* hat sich in den letzten fünf Jahren im Bestreben nach Verbesserung deutlich verändert.

Besonders augenfällig wird in den Beiträgen zu einzelnen Angabebereichen, wie stark sich die streng korpusgestützte lexikografische Arbeit auswirkte:

- Die Korpusbefunde haben in manchen Angabebereichen zu wichtigen Ergänzungen geführt (z. B. im Bereich der sinn- und sachverwandten Wörter, wo neue paradigmatische Relationen beschrieben werden, oder im Bereich

²² Zu *grammis* vgl. <http://hypermedia.ids-mannheim.de/index.html> (Stand: 10.05.2010).

der Gebrauchsbesonderheiten, wo eine neue Angabe die themengebundene Verwendung eines Stichwortes beschreibt). Sie gaben auch den Anstoß dazu, über ursprünglich für die lexikografische Beschreibung zugrunde gelegte, lexikologische Konzepte erneut nachzudenken (z. B. im Bereich der semantischen Umgebung und der lexikalischen Mitspieler).

- Die Arbeit an inzwischen etwa 1.400 beschriebenen Stichwörtern (Stand 2010) wies auch immer wieder auf die Notwendigkeit hin, dass noch offene Fragen der Lemmatisierung (z. B. von Eigennamen) entschieden werden mussten, was zwischenzeitlich geschehen ist. Aufgrund der praktischen Erfahrungen mit dem Korpus musste auch die Datenmodellierung für die Wortartikel an manchen Stellen ergänzt werden, was nur ein Aspekt der hier beschriebenen praktischen Erfahrungen mit der gewählten Modellierung ist.
- Bestimmte Angabebereiche (besonders die Bedeutungserläuterungen, aber auch typische Verwendungsmuster, themengebundene Verwendungen des Stichwortes, grammatische Angaben) sind sehr intensiv mit anderen Angabebereichen im Wortartikel vernetzt. Im Nachdenken über diese Vernetzungen an verschiedenen Stellen in diesem Band wird indirekt deutlich, dass sich im Korpus naturgemäß in den gleichen Belegen – und also nicht auf verschiedene Belege verteilt – zeigt, was die Lexikografen hingegen in verschiedenen Angabebereichen innerhalb eines Wortartikels beschreiben.

Auch dem Medium Internet ist ein großer Einfluss auf die Fortentwicklung von *elexiko* nicht abzuspüren: Die reichen Möglichkeiten der Hypertextualisierung des Wörterbuches, der theoretisch unbegrenzte Platz auf den Bildschirmseiten und die Chance, neue Zugriffsmöglichkeiten auf die Wortartikel zu eröffnen, haben sich in den vergangenen fünf Jahren praktischer Artikelarbeit auf viele Bereiche des Wörterbuches ausgewirkt, nicht zuletzt auf die dichte Vernetzung der lexikografischen Angaben untereinander. In den Beiträgen in diesem Band wird deshalb wiederholt auf Aspekte der Vernetzung sowie auf geplante Erweiterungen der Zugriffsmöglichkeiten hingewiesen. Aber auch die Bereicherung der Wortartikel um multimediale Elemente wie Illustrationen und Ausspracheangaben in Form von Hörbeispielen gehört in diesen Kontext.

Daneben wird deutlich, dass das Projekt *elexiko* zunehmend sprachtechnologische Anwendungen (z. B. Morphologietools wie Morphisto) nutzt, um redaktionell erarbeitete und automatisch erstellte (allerdings meist redaktionell geprüfte) Angaben zu kombinieren. Dies ist vor allem für eine bessere Benutzbarkeit von *elexiko* wichtig, da nur so relativ schnell verlässliche Informationen zu möglichst vielen Stichwörtern angeboten werden können. Eine

ausschließlich redaktionelle Bearbeitung²³ der ca. 300.000 Stichwörter ist im Grunde ohne Rückgriff auf bestehende Wörterbuchressourcen nicht realisierbar, was dem Prinzip der Korpusbasiertheit allerdings fundamental widerspricht. Außerdem würde sich das Publikationstempo von *elexiko* dadurch deutlich verlangsamen und so merklich weniger erfolgreiche Nachschlagehandlungen ermöglichen. *elexiko* wird daher auch weiterhin die korpusgestützte, reflektierte und komplexe Beschreibung einzelner Wortschatzausschnitte mit der (teil-)automatischen, lexikografisch gesteuerten Anreicherung der großen Masse der Stichwörter kombinieren und dabei regelmäßig prüfen, welche neuen sprachtechnologischen Anwendungen weitere qualitativ hochwertige Ergebnisse zur Einbindung in das Wörterbuch versprechen.

6. Literaturverzeichnis

6.1 Wörterbücher

Algemeen Nederlands Woordenboek. Internet: www.inl.nl/en/nieuws/732-algemeen-nederlands-woordenboek-online (Stand: 30.04.2010).

elexiko (2003ff.). In: Institut für Deutsche Sprache (Hg.): OWID – Online-Wortschatz-Informationssystem Deutsch. Mannheim. Internet: www.elexiko.de (Stand: 30.04.2010).

6.2 Forschungsliteratur

Bock, Wolfgang (2008): Technische Aspekte des OWID-Portals. In: Klosa (Hg.), S. 37-44. Internet: www.ids-mannheim.de/pub/laufend/opal/privat/pdf/opal08-1_bock.pdf (Stand: 18.05.2010).

Hahn, Marion/Klosa, Annette/Müller-Spitzer, Carolin/Schnörch, Ulrich/Storjohann, Petra (2008): *elexiko* – das elektronische, lexikografisch-lexikologische korpusbasierte Wortschatzinformationssystem. Zur Neukonzeption, Erweiterung und Revision einzelner Angabebereiche. In: Klosa (Hg.), S. 57-85. Internet: www.ids-mannheim.de/pub/laufend/opal/privat/pdf/opal08-1_hahn-klosa-mueller-spitzer.pdf (Stand: 18.05.2010).

Haß, Ulrike (2005): *elexiko* – Das Projekt. In: Haß (Hg.), S. 1-17.

Haß, Ulrike (Hg.) (2005): Grundfragen der elektronischen Lexikographie. *elexiko* – das Online-Informationssystem zum deutschen Wortschatz. (= Schriften des Instituts für Deutsche Sprache 12). Berlin/New York.

Klosa, Annette (2007): Korpusgestützte Lexikographie: besser, schneller, umfangreicher? In: Kallmeyer, Werner/Zifonun, Gisela (Hg.): Sprachkorpora – Datenmengen und Erkenntnisfortschritt. Jahrbuch 2006 des Instituts für Deutsche

²³ Einen Einblick in die personelle Ausstattung des Projektes geben die *elexiko*-Projektseiten im Internet.

- Sprache. Berlin/New York, S. 105-122. Internet: www.ids-mannheim.de/pub/laufend/jahrbuch/jb2006.html (Stand: 11.05.2010).
- Klosa, Annette (2008): Vorwort. In: Klosa (Hg.), S. 3-4. Internet: www.ids-mannheim.de/pub/laufend/opal/privat/pdf/opal08-1_vorw.pdf (Stand: 11.05.2010).
- Klosa, Annette (Hg.) (2008): Lexikografische Portale im Internet. (= OPAL – Online publizierte Arbeiten zur Linguistik 1/2008). Mannheim. Internet: www.ids-mannheim.de/pub/laufend/opal/privat/opal08-1.html (Stand: 11.05.2010).
- Klosa, Annette (2009): Außentexte in elektronischen Wörterbüchern. In: Beijik, Egbert (Hg.): Fons verborum: feestbundel voor prof. dr. A. M. F. J. (Fons) Moerdijk, aangeboden door vrienden en collegas bij zijn afscheid van het Instituut voor Nederlandse Lexikologie. Amsterdam, S. 49-60.
- Klosa, Annette (2010): On the combination of automated information and lexicographically interpreted information in two German online dictionaries. In: Granger, Sylviane/Paquot, Magali (Hg.): eLexicography in the 21st century: new challenges, new applications. Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009. (= Cahier du Cental 7). Louvain-la-Neuve, S. 157-163.
- Meyer, Peter/Müller-Spitzer, Carolin (2010): Consistency of sense relations in a lexicographic context. In: Barbu Mititelu, Verginica/Pekar, Viktor/Barbu, Eduard (Hg.): Proceedings of the Workshop „Semantic Relations. Theory and Applications“, 18 May 2010, at the International Conference on Language Resources and Evaluation (LREC), 2010, Malta. Internet: <http://www.lrec-conf.org/proceedings/lrec2010/workshops/W9.pdf> (Stand: 08.10.2010).
- Müller-Spitzer, Carolin (2008): The lexicographic portal of the IDS. Connecting heterogeneous lexicographic resources by a consistent concept of data modelling. In: Bernal, Elisenda/DeCesaris, Janet (Hg.): Proceedings of the Thirteenth EURALEX International Congress, Barcelona, Spain, July 15th-19th, 2008. (= Série Activitats 20). Barcelona, S. 457-461.
- Müller-Spitzer, Carolin/Schneider, Roman (2009): Ein XML-basiertes Datenbanksystem für digitale Wörterbücher – Ein Werkstattbericht aus dem Institut für Deutsche Sprache. In: it-Information Technology 51, 4, S. 197-206.
- Storjohann, Petra (2005a): Corpus-driven vs. corpus-based approach to the study of relational patterns. In: Proceedings of the Corpus Linguistics Conference 2005 in Birmingham. Internet: www.corpus.bham.ac.uk/PCLC/. Stand: 30.04.2010.
- Storjohann, Petra (2005b): Das *lexiko*-Korpus: Aufbau und Zusammensetzung. In: Haß (Hg.), S. 55-70.

6.3 Internetressourcen

- Belica, Cyril (1995): Statistische Kollokationsanalyse und -clustering. Korpuslinguistische Analysemethode. Mannheim. Internet: <http://corpora.ids-mannheim.de/ccdb/> (Stand: 30.04.2010).

- Belica, Cyril (2001-2007): Kookkurrenzdatenbank CCDB – V3.2. Eine korpuslinguistische Denk- und Experimentierplattform für die Erforschung und theoretische Begründung von systemisch-strukturellen Eigenschaften von Kohäsionsrelationen zwischen den Konstituenten des Sprachgebrauchs. Mannheim. Internet: <http://corpora.ids-mannheim.de/ccdb/> (Stand: 30.04.2010).
- BZ*Velexiko* – Benutzeradaptive Zugänge und Vernetzung in *elexiko*: Internet: www.ids-mannheim.de/lexik/BZVelexiko (Stand: 30.04.2010).
- canoo.net – Deutsche Wörterbücher und Grammatik. Internet: www.canoo.net (Stand: 30.04.2010).
- CCDB – Kookkurrenzdatenbank – V3.2. Internet: <http://corpora.ids-mannheim.de/ccdb/> (Stand: 30.04.2010).
- COSMAS II. Internet: <https://cosmas2.ids-mannheim.de/cosmas2-web/> (Stand: 30.04.2010).
- DEREKO – Deutsches Referenzkorpus des IDS. Internet: www.ids-mannheim.de/kl/projekte/korpora/ (Stand: 30.04.2010).
- EDAS – Electronic Dictionary Administration System / Lexikographisches Redaktions- und Recherchesystem für digitale Wörterbücher. Internet: www.ids-mannheim.de/gra/texttechnologie.html (Stand: 30.04.2010).
- grammis – das grammatische Informationssystem des Instituts für Deutsche Sprache. Internet: <http://hypermedia.ids-mannheim.de/index.html> (Stand: 30.04.2010).
- Morphisto. Internet: www.ids-mannheim.de/ll/TextGrid/morphisto.html (Stand: 30.04.2010).
- OWID – Online-Wortschatz-Informationssystem Deutsch (2008ff.). Institut für Deutsche Sprache (Hg.). Mannheim. Internet: www.owid.de (Stand: 30.04.2010).